

# On the reservoir technique convergence for nonlinear hyperbolic conservation laws – Part I

Stéphane Labbé<sup>a</sup>, Emmanuel Lorin<sup>b,c,\*</sup>

<sup>a</sup> Département de Mathématiques, Université Joseph Fourier, 38041 Grenoble Cedex 9, France

<sup>b</sup> Faculty of Science, University of Ontario Institute of Technology, Oshawa, ON, L1H 7K4, Canada

<sup>c</sup> Centre de Recherches Mathématiques, Montréal, Québec, H3T 1J4, Canada

## ARTICLE INFO

### Article history:

Received 22 January 2008

Available online 5 March 2009

Submitted by J. Guermont

### Keywords:

Finite volume scheme

Numerical diffusion

Convergence

Hyperbolic systems

## ABSTRACT

This paper is devoted to the analysis of flux schemes coupled with the reservoir technique for approximating hyperbolic equations and linear hyperbolic systems of conservation laws [F. Alouges, F. De Vuyst, G. Le Coq, E. Lorin, The reservoir scheme for systems of conservation laws, in: *Finite Volumes for Complex Applications, III*, Porquerolles, 2002, Lab. Anal. Topol. Probab. CNRS, Marseille, 2002, pp. 247–254 (electronic); F. Alouges, F. De Vuyst, G. Le Coq, E. Lorin, Un procédé de réduction de la diffusion numérique des schémas à différence de flux d'ordre un pour les systèmes hyperboliques non linéaires, *C. R. Math. Acad. Sci. Paris* 335 (7) (2002) 627–632; F. Alouges, F. De Vuyst, G. Le Coq, E. Lorin, The reservoir technique: A way to make Godunov-type schemes zero or very low diffusive. Application to Colella–Glaz, *Eur. J. Mech. B Fluids* 27 (6) (2008)]. We prove the long time convergence of the reservoir technique and its TVD property for some specific but still general configurations. Proofs are based on a precise study of the treatment by the reservoir technique of shock and rarefaction waves.

© 2009 Elsevier Inc. All rights reserved.

## 1. Introduction

### 1.1. Generalities

This paper is devoted to the convergence of the reservoir technique introduced in [1–3]. This method allows us to avoid or to reduce drastically the numerical diffusion of flux schemes used to compute solutions to hyperbolic systems of conservation laws. This technique is based on the fact that for a scalar linear advection equation, a flux scheme [8] is stable and non-diffusive with the CFL (Courant–Friedrich–Lewy condition) equal to 1. Let us indeed consider the advection problem, with  $a > 0$ :  $\partial_t v + a \partial_x v = 0$ ,  $x \in \mathbb{R}$ ,  $t \geq 0$ , and a conservative explicit upwind scheme approximating this equation:  $\forall n \geq 0, \forall j \in \mathbb{Z}, u_j^{n+1} = u_j^n - a \Delta t / \Delta x (u_j^n - u_{j-1}^n)$  where  $u_j^n$  is an approximation of  $v$  at  $(j \Delta x, n \Delta t)$ . As it is well known this scheme is stable under the CFL criterion:  $\lambda = a \frac{\Delta t}{\Delta x} \in ]0, 1]$ , and that taking  $\Delta t = \Delta x / a$  ( $\lambda = 1$ ) leads to  $u_j^{n+1} = u_{j-1}^n$  which corresponds to the exact propagation of the solution at the discrete level. However taking  $\Delta t = \Delta x / a$  is very restrictive (if  $\Delta x$  or  $a$  is not constant in space for example) and we must constraint our scheme to behave as if  $\lambda = 1$  although in general the time step is smaller than  $\Delta t_* := \Delta x / a$ . This is precisely the goal of the reservoir technique. To understand more clearly the principle, we consider the case where  $\Delta t = \Delta t_* / k$  with  $k \in \mathbb{N}^*$ . The idea consists of waiting  $k$  time steps before updating the solution. More generally, let us take several different time steps  $\Delta t_i$  (not necessarily equal), and introduce  $R_j$  a (scalar)

\* Corresponding author at: Faculty of Science, University of Ontario Institute of Technology, Oshawa, ON, L1H 7K4, Canada.

E-mail address: lorin@crm.umontreal.ca (E. Lorin).

reservoir associated to the  $j$ th cell (initialized to 0) and  $c_j$  a positive real CFL counter (also initialized to 0). At each time step, we fill up  $R_j$  with the local current numerical flux difference  $R_j \leftarrow R_j - a\Delta t_i/\Delta x(u_j - u_{j-1})$  and we increment the CFL counter  $c_j \leftarrow c_j + a\Delta t_i/\Delta x$ . Once  $c_j$  reaches 1, we update the solution and reinitialize to zero both the reservoir and the counter.

$$u_j \leftarrow u_j + R_j, \quad c_j \leftarrow 0, \quad R_j \leftarrow 0.$$

In the advection equation case we recover the CFL = 1. Indeed the reservoir technique for one-dimensional linear equation behaves as:

$$u_j^{n+1} = u_j^n - a \sum_{i=1}^k \frac{\Delta t_i}{\Delta x} (u_j^n - u_{j-1}^n) = u_{j-1}^n.$$

In this formula  $n$  denotes the real time step and  $i$  denotes the local iteration index between two time steps such that:  $a \sum_{i=1}^k \Delta t_i/\Delta x = 1$ .

Extending this idea to hyperbolic systems of conservation laws we only update the numerical solution for each discretization point where the local CFL counter reaches 1 or in fact any prescribed value between  $0^+$  and  $1^-$ . More precisely, at each space point and for each characteristic field we attach a counter that measures the local CFL number and a vectorial reservoir that stores the numerical flux. This technique has been successfully applied to nonlinear hyperbolic systems of conservation laws [2,3] and is currently developed for multidimensional equations, see [4]. The reservoir technique is not linked to a particular scheme but is applicable to many kinds of flux schemes [8] based on approximate linear or nonlinear Riemann solvers as Roe [13], VFC [9], Colella and Glaz [6] or exact Riemann solvers as Godunov [11]. Nevertheless, it is when coupled with the Colella–Glaz scheme [6] that the reservoir technique gives the most impressive results [3] compared to higher order classical schemes as ENO [14], WAF [5] for example, as we can see in [12]. In this paper we prove the convergence of the reservoir technique coupled with flux schemes [8] approximating one-dimensional nonlinear scalar hyperbolic conservation laws and linear hyperbolic systems of conservation laws for a general class of initial data.

We are able in this work to prove the long time convergence of the numerical solution obtained via the reservoir technique for shock waves. The reservoir technique main specificity comes from the fact that as for Glimm's scheme, the local error in space, is not accumulated at all times, but can be totally (shock waves) or partially (rarefaction waves) canceled after some iterations. With the same principle, the TVD property is proven. Under some arithmetical conditions on the characteristic velocities, we can even prove that the reservoir technique captures the exact solution at the discrete level.

The paper is organized as follows. In Section 1.2 we recall some important features of flux schemes and we set the studied problem. In Section 2 we prove the convergence and the TVD property of the reservoir technique coupled with flux schemes for nonlinear hyperbolic equations of conservation laws. Then the case of linear systems is studied in Section 2.3. The proof of convergence for nonlinear systems is much more technical and will be presented in a forthcoming paper. Finally, we conclude in Section 3.

## 1.2. Flux schemes and reservoir technique

We consider first order finite volume schemes based on Riemann solvers [8] for solving nonlinear hyperbolic conservation laws:

$$\begin{cases} v_t + f(v)_x = 0, & x \in \mathbb{R}, t \geq 0, \\ v(x, 0) = v_0(x), & x \in \mathbb{R}. \end{cases}$$

**Definition 1.1.** An explicit conservative three-point finite volume scheme with Riemann solver writes:

$$u_j^{n+1} = u_j^n - \frac{\Delta t_n}{\Delta x_j} (g(u_j^n, u_{j+1}^n) - g(u_{j-1}^n, u_j^n)), \quad j \in \mathbb{Z}, n \geq 0, \quad (1)$$

where  $u_j^n$  is an approximation to  $\frac{1}{\Delta x_j} \int_{x_{j-1/2}}^{x_{j+1/2}} v(x, t_n) dx$  and  $\Delta t_n$  is the time step at iteration  $n$  (verifying a CFL condition). The interfacial numerical flux is denoted by  $g(u, v)$  and is computed with an exact or approximate Riemann solver.

Among these finite volume schemes we will consider in this paper finite volume flux schemes:

**Definition 1.2.** A flux scheme is a finite volume scheme that writes:

$$u_j^{n+1} = u_j^n - \frac{\Delta t_n}{\Delta x_j} (g(u_j^n, u_{j+1}^n) - g(u_{j-1}^n, u_j^n)),$$

where  $g$  is the numerical flux given by:

$$g(s, t) = \frac{1}{2} (f(s) + f(t)) - \frac{\Lambda(s, t)}{2} (f(t) - f(s)), \quad (2)$$

and  $\Lambda = P \operatorname{diag}(\operatorname{sgn}(\lambda_i)) P^{-1}$  is a sign matrix, where  $(\lambda_i)_i$  are the eigenvalues of  $df$ , and  $P$  is the matrix composed of the eigenvectors of  $df$ . See [8,9] for additional details.

Recall that explicit flux schemes are first order convergent schemes under a classical CFL stability condition [8]. In the reservoir technique framework, we consider two kinds of flux schemes: those with a numerical flux computed using the Colella–Glaz Riemann solver and those using a linearized Riemann solver (VFFC or Roe for instance). The main difference between these two approaches comes from the fact that, the Colella–Glaz scheme consists of solving Riemann problems at the mesh interfaces using a nonlinear solver and an original flux difference decomposition, in opposite with linear flux schemes that use linearized Riemann solvers.

Note that in the sequel the space step will be constant:  $\Delta x_j = \Delta x$  for all  $j \in \mathbb{Z}$ . We will denote by  $\Delta t$  the maximum of all  $\Delta t_i$ , where  $\Delta t_i$  is the time step at time iteration  $i$ :

$$\Delta t = \max_{i \in \mathbb{N}} \Delta t_i. \quad (3)$$

**Main goal:** Consider an initial data  $u^0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \cap BV(\mathbb{R})$ . As usual we will consider weak discontinuities, that is  $|u_{\text{left state}} - u_{\text{right state}}|$  is small enough (even if this assumption is only necessary for the rarefaction wave study). Since the time and space steps are linked by a stability condition (typically by a positive constant  $a$  such that  $\Delta t / \Delta x \leq a$ ), we could expect at most that there exists  $c > 0$  such that

$$\|v(\cdot, t_n) - u^n\|_{L^1} \leq \|v(\cdot, 0) - u^0\|_{L^1} + c t_n \Delta t, \quad \forall n \in \mathbb{N}.$$

In fact we can prove for instance for solutions with shock waves only, the existence of a constant  $c$  independent of  $t_n$  such that:

$$\|v(\cdot, t_n) - u^n\|_{L^1} \leq \|v(\cdot, 0) - u^0\|_{L^1} + c \Delta t, \quad \forall n \in \mathbb{N},$$

where  $v$  denotes the exact solution of the continuous problem and  $u^n$  its reservoir approximation. Note that this property is very close to Lagoutière and Desprès' estimate with their antidiffusive numerical scheme presented in [7]. This allows us to prove a long-time convergence at least for solutions with shock waves only. This is a powerful property that very few finite volume schemes satisfy.

Suppose that at time  $t = 0$  the initial data is given by  $u^0$ . Then by definition of the reservoir technique, for all  $j \in \mathbb{Z}$  there exists  $p_j$  in  $\mathbb{N}$  such that  $c_j^{p_j-1} > 0$  and  $c_j^{p_j} = 0$  with:

$$\begin{cases} u_j^l = u_j^0, & 0 \leq l < p_j, \\ u_j^{p_j} = u_j^0 - \sum_{l=0}^{p_j-1} \frac{\Delta t_l}{\Delta x} (g(u_j^0, u_{j+1}^0) - g(u_{j-1}^0, u_j^0)), & j \in \mathbb{Z}. \end{cases}$$

That implies that the reservoir technique combined with a flux scheme based on exact or approximate Riemann solvers consists of choosing a particular local time step which is equal to  $\sum_{l=0}^{p_j-1} \Delta t_l$ .

For a classical flux scheme the solution  $w$  after  $p_j$  iterations in the cell  $j \in \mathbb{Z}$  (with  $w^0 = u^0$ ) from 0 to  $t_{p_j}$  is naturally given by:

$$\begin{cases} w_j^1 = u_j^0 - \frac{\Delta t_0}{\Delta x} (g(u_j^0, u_{j+1}^0) - g(u_{j-1}^0, u_j^0)), \\ w_j^l = w_j^{l-1} - \frac{\Delta t_{l-1}}{\Delta x} (g(w_j^{l-1}, w_{j+1}^{l-1}) - g(w_{j-1}^{l-1}, w_j^{l-1})), & \forall l \leq p_j. \end{cases}$$

That is:

$$w_j^{p_j} = u_j^0 - \sum_{l=0}^{p_j-1} \frac{\Delta t_l}{\Delta x} (g(w_j^l, w_{j+1}^l) - g(w_{j-1}^l, w_j^l)), \quad j \in \mathbb{Z}.$$

For a first order finite volume scheme approximating a system of conservation laws, let us recall there exists  $c > 0$ :

$$\|v(\cdot, t_n) - w^n\|_{L^1} \leq \|v(\cdot, 0) - w^0\|_{L^1} + c t_n \Delta t, \quad \forall n \in \mathbb{N}.$$

Unfortunately, error estimates for the reservoir technique combined with a flux scheme cannot be deduced easily from the error estimate for the considered flux scheme. Indeed using the same notations as above we have after  $l < p_j$  iterations:

$$|w_j^l - u_j^l| = \left| \sum_{k=0}^{l-1} \frac{\Delta t_k}{\Delta x} (g(w_j^k, w_{j+1}^k) - g(w_{j-1}^k, w_j^k)) \right|. \quad (4)$$

We can deduce that there exists a positive constant  $d$  such that the error  $\|w^l - u^l\|_{L^1}$  is always strictly less the  $d \Delta t$  (recall that  $\Delta t$  and  $\Delta x$  are linked by a stability condition).

## 2. Scalar equation

### 2.1. Numerical scheme

We consider a nonlinear scalar conservation law

$$\begin{cases} v_t + f(v)_x = 0, & x \in \mathbb{R}, t \geq 0, \\ v(\cdot, 0) = v_0(\cdot) \in BV(\mathbb{R}) \cap L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}), & x \in \mathbb{R}, \end{cases} \quad (5)$$

with  $f$  a regular strictly convex function. As usual we search for an approximation  $u_j^n$  of the solution  $v$ . The initial data is given by:  $(u_j^0)_j = (\frac{1}{\Delta x} \int_{x_{j-1/2}}^{x_{j+1/2}} v_0(x) dx)_j$ . In this case, the classical upwind scheme writes:

$$u_j^{n+1} = u_j^n - \frac{\Delta t_n}{\Delta x} (f_{j+\frac{1}{2}}^n - f_{j-\frac{1}{2}}^n), \quad j \in \mathbb{Z}, n \geq 1,$$

where  $f_{j-\frac{1}{2}}^n$  is the interfacial upwinded flux:

$$f_{j-\frac{1}{2}}^n = \frac{f(u_j^n) + f(u_{j-1}^n)}{2} - \frac{1}{2} \operatorname{sgn}(f'(u_{j-1/2}^n))(f(u_j^n) - f(u_{j-1}^n)),$$

with  $\Delta t_n = t_{n+1} - t_n$  and  $u_{j-1/2}^n$  being the approximate value of the solution at the interface. In the framework of one-dimensional nonlinear scalar equations two kinds of waves can appear: rarefaction and shock waves. At each interface  $j - 1/2$  between cells  $j - 1$  and  $j$ , the Lax entropy condition states that  $f'(u_{j-1}^n) > f'(u_j^n)$  generates an entropy shock wave whereas  $f'(u_j^n) > f'(u_{j-1}^n)$  generates a rarefaction wave. Let us also recall that from the Rankine–Hugoniot condition, a shock wave at the interface  $j - 1/2$  has a speed equal to:

$$\sigma_{j-1/2}^n = \frac{f(u_j^n) - f(u_{j-1}^n)}{u_j^n - u_{j-1}^n}.$$

Using this interfacial speed, we propose a method to update reservoirs and counters, upwinding the interfacial numerical flux. Let us denote by  $\lambda_-^n = f'(u_{j-1}^n)$  and by  $\lambda_+^n = f'(u_j^n)$  the left and right characteristic speeds. The idea is to determine the time step according to the largest speed in the wave. For shock waves, it is given by the Rankine–Hugoniot condition, whereas for rarefaction fans, it is given by the largest speed  $\lambda_-^n$  or  $\lambda_+^n$ . Moreover, in the case of sonic points ( $\lambda_-^n \leq 0 \leq \lambda_+^n$ ), we must split the wave in two parts, one going to the left and the other one going to the right. Initially, we set to zero all the reservoirs and counters  $c_j^0 = R_j^0 = 0$  for all  $j$ .

– Right shock wave ( $\lambda_+^n \leq \lambda_-^n$  and  $\sigma_{j-1/2}^n > 0$ )

$$\begin{pmatrix} u_j^{n+1} \\ c_j^{n+1} \\ R_j^{n+1} \end{pmatrix} = \begin{cases} (u_j^n, c_j^n + \frac{\sigma_{j-1/2}^n \Delta t_n}{\Delta x}, R_j^n - \frac{\Delta t_n}{\Delta x} (f(u_j^n) - f(u_{j-1}^n)))^T, & \text{if } c_j^n + \frac{\sigma_{j-1/2}^n \Delta t_n}{\Delta x} < 1, \\ (u_j^n + R_j^n - \frac{\Delta t_n}{\Delta x} (f(u_j^n) - f(u_{j-1}^n)), 0, 0)^T, & \text{if } c_j^n + \frac{\sigma_{j-1/2}^n \Delta t_n}{\Delta x} = 1. \end{cases} \quad (6)$$

– Left shock wave ( $\lambda_+^n \leq \lambda_-^n$  and  $\sigma_{j-1/2}^n < 0$ )

$$\begin{pmatrix} u_j^{n+1} \\ c_j^{n+1} \\ R_j^{n+1} \end{pmatrix} = \begin{cases} (u_j^n, c_j^n + \frac{|\sigma_{j-1/2}^n| \Delta t_n}{\Delta x}, R_j^n - \frac{\Delta t_n}{\Delta x} (f(u_{j+1}^n) - f(u_j^n)))^T, & \text{if } c_j^n + \frac{|\sigma_{j-1/2}^n| \Delta t_n}{\Delta x} < 1, \\ (u_j^n + R_j^n - \frac{\Delta t_n}{\Delta x} (f(u_{j+1}^n) - f(u_j^n)), 0, 0)^T, & \text{if } c_j^n + \frac{|\sigma_{j-1/2}^n| \Delta t_n}{\Delta x} = 1. \end{cases}$$

– Right rarefaction wave ( $\lambda_+^n \geq \lambda_-^n \geq 0$ )

$$\begin{pmatrix} u_j^{n+1} \\ c_j^{n+1} \\ R_j^{n+1} \end{pmatrix} = \begin{cases} (u_j^n, c_j^n + \frac{\lambda_+^n \Delta t_n}{\Delta x}, R_j^n - \frac{\Delta t_n}{\Delta x} (f(u_j^n) - f(u_{j-1}^n)))^T, & \text{if } c_j^n + \frac{\lambda_+^n \Delta t_n}{\Delta x} < 1, \\ (u_j^n + R_j^n - \frac{\Delta t_n}{\Delta x} (f(u_j^n) - f(u_{j-1}^n)), 0, 0)^T, & \text{if } c_j^n + \frac{\lambda_+^n \Delta t_n}{\Delta x} = 1. \end{cases}$$

– Left rarefaction wave ( $0 \geq \lambda_+^n \geq \lambda_-^n$ )

$$\begin{pmatrix} u_j^{n+1} \\ c_j^{n+1} \\ R_j^{n+1} \end{pmatrix} = \begin{cases} (u_j^n, c_j^n + \frac{|\lambda_-^n| \Delta t_n}{\Delta x}, R_j^n - \frac{\Delta t_n}{\Delta x} (f(u_{j+1}^n) - f(u_j^n)))^T, & \text{if } c_j^n + \frac{|\lambda_-^n| \Delta t_n}{\Delta x} < 1, \\ (u_j^n + R_j^n - \frac{\Delta t_n}{\Delta x} (f(u_{j+1}^n) - f(u_j^n)), 0, 0)^T, & \text{if } c_j^n + \frac{|\lambda_-^n| \Delta t_n}{\Delta x} = 1. \end{cases}$$

For sonic points (that are not considered in this paper) the scheme is described in [3].

## 2.2. Numerical analysis

In this section, we study the convergence of the reservoir technique for the nonlinear scalar equation (7) with strictly convex flux  $f \in C^3(\mathbb{R})$ . We first consider solutions without wave interactions, which is reasonable for a large class of initial data and at least for a finite time  $T > 0$ . Recall that

$$v_t + f(v)_x = 0, \quad f \in C^3(\mathbb{R}) \text{ and strictly convex, } v_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \cap BV(\mathbb{R}). \quad (7)$$

The following analysis is valid for all discrete initial data  $(u_j^0)_{j \in \mathbb{Z}}$  defined as:

$$u^0 = \sum_{j \in \mathbb{Z}} u_j^0 \mathbf{1}_{[(j-1/2)\Delta x, (j+1/2)\Delta x[},$$

where  $\mathbf{1}$  is equal to 1 on  $[(j-1/2)\Delta x, (j+1/2)\Delta x[$  and equal to zero elsewhere. We consider step-like initial data, such that for some  $j_0 \in \mathbb{Z}$ :

$$u^0 = \sum_{j \leq j_0} u_l \mathbf{1}_{[(j-1/2)\Delta x, (j+1/2)\Delta x[} + \sum_{j > j_0} u_r \mathbf{1}_{[(j-1/2)\Delta x, (j+1/2)\Delta x[}.$$

We first state that for solutions containing only shock waves:

**Theorem 2.1.** *The reservoir scheme solution  $(u^n)$ , approximating a 1d scalar hyperbolic equation of conservation laws with a strictly convex flux (5), with only shock waves, is convergent to the exact solution  $v$ . Moreover, supposing that no wave interaction has occurred for all  $t \leq t_n$ , there exists  $c(t_n) > 0$  ( $c$  depends on  $t_n$ ) such that:*

$$\|u^n - v(\cdot, t_n)\|_{L^1} \leq \|u^0 - v(\cdot, 0)\|_{L^1} + c(t_n)TV(u^0)\Delta t$$

and there exists  $\bar{c} > 0$  such that  $\sup_{n \in \mathbb{N}} c(t_n) \leq \bar{c}$ .

That is, we have a long time convergence for solutions having only shock waves. In fact this result can be extended to solutions containing shock and rarefaction waves. However, we do not have anymore the long time convergence, even if it leads to a better approximation for long times than order 1 schemes.

**Theorem 2.2.** *The reservoir scheme solution  $(u^n)$ , approximating a 1d scalar hyperbolic equation of conservation laws with a strictly convex flux (5) is convergent to the exact solution  $v$ . Moreover, supposing that no wave interaction has occurred for all  $t \leq t_n$ , there exist  $c(t_n) > 0$  ( $c$  depends on  $t_n$ ),  $K(n) \in \mathbb{N}^*$ , and  $d = d(TV(u^0)) > 0$  ( $d$  depends on  $TV(u_0)$ ) such that:*

$$\|u^n - v(\cdot, t_n)\|_{L^1} \leq \|u^0 - v(\cdot, 0)\|_{L^1} + c(t_n)TV(u^0)\Delta t + dK(n)\Delta t^2$$

and there exists  $\bar{c} > 0$  such that  $\sup_{n \in \mathbb{N}} c(t_n) \leq \bar{c}$  and  $K(n) < n$ .

Although this result seems a little disappointing, it has to be noticed that  $K(n)$  tends to infinity less rapidly than  $n$  (see next section).

**Proof of Theorem 2.2.** As there exist two kinds of non-trivial waves for such equations (shock and rarefaction waves) we propose to study the two situations and we split the proof in two parts.

### Step 1. Shock waves.

Consider a time iteration  $n_0$  such that all the counters are equal to zero (ex:  $n_0 = 0$ ). For all  $n \in \mathbb{N}^*$  and  $j \in \mathbb{Z}$ , while  $c_j^{n_0+n} \neq 0$  by the definition of the reservoir technique, we freeze the physical flux associated to the cell  $j$  at time  $t_{n_0}$ . We will study (by convention) the case of a right-entropy shock with a discontinuity located at  $x_{j_0-1/2}$  at time  $t_{n_0}$  and of velocity  $\sigma_{j_0-1/2}^{n_0}$ . As  $\text{sgn}(f')$  is positive, we update the solution as follows (for  $n = 1$ ):

$$u_j^{n_0+1} = u_j^{n_0} - \frac{\Delta t_{n_0}}{\Delta x} (f(u_j^{n_0}) - f(u_{j-1}^{n_0})) \mathbf{1}_{c_j^{n_0+1}=0}.$$

More generally, for every  $n \in \mathbb{N}^*$  and  $j$  in  $\mathbb{Z}$  such that  $c_j^{n_0+n} > 0$  one has

$$\begin{aligned} u_j^{n_0+n} &= u_j^{n_0+n-1} - \frac{\Delta t_{n_0+n-1}}{\Delta x} (f(u_j^{n_0}) - f(u_{j-1}^{n_0})) \mathbf{1}_{c_j^{n_0+n}=0} \\ &= u_j^{n_0} - \sum_{i=0}^{n-1} \frac{\Delta t_{n_0+i}}{\Delta x} (f(u_j^{n_0}) - f(u_{j-1}^{n_0})) \mathbf{1}_{c_j^{n_0+n}=0}. \end{aligned} \quad (8)$$

Recall that the time step is chosen such that:

$$\Delta t_n = \min_j \left( (1 - c_j^n) \frac{\Delta x}{|\sigma_{j-1/2}^n|} \right).$$

We can then prove:

**Lemma 2.1.** *Let us suppose that the exact solution  $v$  is a shock wave with a front located at  $(j_0 - 1/2)\Delta x$  at time  $t_{n_0}$ . Then there exist  $j' \in \mathbb{Z}$ ,  $j' = j_0 + \text{sgn}(\sigma_{j_0-1/2}^{n_0})$  and  $n' > n_0$  such that*

$$|v(x_{j'}, t_{n'}) - u_{j'}^{n'}| = |v(x_{j_0}, t_{n_0}) - u_{j_0}^{n_0}|,$$

with

$$|v(x_{j'}, t_k) - u_{j'}^k| = |v(x_{j_0}, t_{n_0}) - u_{j_0}^{n_0}| + \mathcal{O}(\Delta t), \quad n_0 \leq k < n'.$$

**Proof.** The shock is a right-entropy shock. As  $f'$  is increasing, while  $c_{j_0}^{n_0+n} \neq 0$  one has  $c_{j_0-1}^{n_0+n} \neq 0$ . So that the corresponding physical flux  $f(u_{j_0-1}^{n_0})$  is not updated. Recall also that  $f(u_{j_0}^{n_0}) - f(u_{j_0-1}^{n_0}) = \sigma_{j_0-1/2}^{n_0}(u_{j_0}^{n_0} - u_{j_0-1}^{n_0})$  by the Rankine–Hugoniot jump condition.

Let us now denote by  $k_{j_0}$  the index such that:  $c_{j_0}^{n_0+k_{j_0}} = 0$ , that is  $\sum_{l=0}^{k_{j_0}-1} \frac{\Delta t_{n_0+l}}{\Delta x} \sigma_{j_0-1/2}^{n_0} = 1$ . Then

$$u_{j_0}^{n_0+k_{j_0}} = u_{j_0}^{n_0} - \sum_{l=0}^{k_{j_0}-1} \frac{\Delta t_{n_0+l}}{\Delta x} \sigma_{j_0-1/2}^{n_0} (u_{j_0}^{n_0} - u_{j_0-1}^{n_0}) = u_{j_0-1}^{n_0}. \quad (9)$$

At time  $t_{n_0+k_{j_0}}$  the exact solution is obviously given by:

$$v(x_{j_0}, t_{n_0+k_{j_0}}) = v(x_{j_0}, t_{n_0}). \quad (10)$$

Because of (9) and (10) and as we have consider a single shock wave solution, we obtain

$$\sum_{j \in \mathbb{Z}} |v(x_j, t_{n_0+k_{j_0}}) - u_j^{n_0+k_{j_0}}| = \sum_{j \in \mathbb{Z}} |v(x_j, t_{n_0}) - u_j^{n_0}|.$$

Then, in the current cell and at time  $t_{n_0+k_{j_0}}$  the reservoir technique combined with the flux scheme has captured the exact shock solution. By induction and symmetry the result is also valid for left entropy-shocks.  $\square$

More generally, for all  $j \in \mathbb{Z}$  and for all time  $t_{n_0:j_i}$  with  $i \in \mathbb{N}$ , defined such that  $c_j^{n_0:j_i} = 0$ , the numerical solution will be exact in the cell  $j$ . That is, at times  $t_{n_0:j_i}$  we have:

$$|v(x_j, t_{n_0:j_i}) - u_j^{n_0:j_i}| = |v(x_j, t_{n_0}) - u_j^{n_0}|.$$

However, between the times  $t_{n_0:j_i}$  and  $t_{n_0:j_{i+1}}$  because of (4), we have the following estimate:

$$|v(x_j, t_k) - u_j^k| = |v(x_j, t_{n_0}) - u_j^{n_0}| + \mathcal{O}((u_j^{n_0} - u_{j-1}^{n_0})\Delta t), \quad \forall k \in \{n_0:j_i + 1, \dots, n_0:j_{i+1} - 1\}, \quad (11)$$

as the solution in the cell  $j$  has been frozen until the local CFL number reaches 1 (and the counter is set to zero). These results allow to prove that the global error remains at order 1 but in each cell there exist times for which the error is zero: then using that  $u^0$  belongs to  $BV(\mathbb{R})$  and Lemma 2.1, for shock waves, there exists a real constant  $c_1 > 0$  such that:

$$\|v(\cdot, t_n) - u^n\|_{L^1} \leq \|v(\cdot, 0) - u^0\|_{L^1} + c_1 TV(u^0) \Delta t, \quad \forall n \in \mathbb{N}. \quad (12)$$

## Step 2. Rarefaction waves.

We now suppose that the exact solution of the continuous problem is a right-rarefaction wave.

Then for all  $n \in \mathbb{N}$ , and  $j \in \mathbb{Z}$ ,  $\lambda_j^n \leq \lambda_{j+1}^n$  (entropy condition), where  $\lambda_j^n = f'(u_j^n)$ .

Consider  $n_0$  such that the  $j$ th counter is equal to zero (ex:  $n_0 = 0$ ). Then for  $n \in \mathbb{N}$ , while  $c_j^{n_0+n} > 0$  and by the definition of the reservoir scheme:

$$u_j^{n_0+n} = u_j^{n_0} \quad \text{and} \quad R_j^{n_0+n} = R_j^{n_0+n-1} - \frac{\Delta t_{n_0+n-1}}{\Delta x} (f(u_j^{n_0+n-1}) - f(u_{j-1}^{n_0+n-1})).$$

But as  $c_j^{n_0+n-1} \neq 0$  then  $u_j^{n_0+n-1} = u_j^{n_0}$  and there exists  $k_j$  such that

$$u_j^{n_0+k_j} = u_j^{n_0} - \sum_{l=0}^{k_j-1} \frac{\Delta t_{n_0+l}}{\Delta x} (f(u_j^{n_0}) - f(u_{j-1}^{n_0+l})).$$

It has to be noticed that even if  $\lambda_{j-1}^n \leq \lambda_j^n$  between  $t_{n_0}$  and  $t_{n_0+k_j}$ , the solution at  $j-1$  could have been updated (excepted if  $c_{j-1}^{n_0} = 0$ ). However  $\lambda_{j-1}^n \leq \lambda_j^n$  implies necessarily that the solution in the cell  $j-1$  has been updated at most one time. Then:

$$u_j^{n_0+k_j} = u_j^{n_0} - \sum_{l=0}^{k_j-1} \frac{\Delta t_{n_0+l}}{\Delta x} (f(u_j^{n_0}) - f(u_{j-1}^{n_0+l}))$$

with

$$u_{j-1}^{n_0+k_j-1} = \begin{cases} u_{j-1}^{n_0}, & \text{if } c_{j-1}^{n_0+k} \neq 0, \forall k \in \{1, \dots, k_j-1\}, \\ u_{j-1}^{n_0} + \mathcal{O}(\Delta t), & \text{if } \exists k \in \{1, \dots, k_j-1\} / c_{j-1}^{n_0+k} = 0. \end{cases}$$

For weak enough initial data  $|u_j^0 - u_{j-1}^0| = \mathcal{O}(\Delta x)$ , for all  $j$  in  $\mathbb{Z}$ , and by regularity of  $f$  we then have:

$$u_j^{n_0+k_j} = u_j^{n_0} - \sum_{l=0}^{k_j-1} \frac{\Delta t_{n_0+l}}{\Delta x} \left( \lambda_j^{n_0} (u_j^{n_0} - u_{j-1}^{n_0}) + \frac{f''(u_j^{n_0})}{2} (u_j^{n_0} - u_{j-1}^{n_0})^2 + \mathcal{O}((u_j^{n_0} - u_{j-1}^{n_0}) \Delta t) + \mathcal{O}((u_j^{n_0} - u_{j-1}^{n_0})^3) \right).$$

That is, using that  $c_j^{n_0+k_j-1} = 0$  or

$$\sum_{l=0}^{k_j-1} \frac{\Delta t_{n_0+l}}{\Delta x} f'(u_j^{n_0}) = 1 \quad (13)$$

we obtain from (13):

$$u_j^{n_0+k_j} = u_{j-1}^{n_0} - \alpha_j^{n_0} \frac{f''(u_j^{n_0})}{2} (u_{j-1}^{n_0} - u_j^{n_0})^2 + \mathcal{O}((u_j^{n_0} - u_{j-1}^{n_0}) \Delta t) + \mathcal{O}((u_j^{n_0} - u_{j-1}^{n_0})^3), \quad (14)$$

with

$$\alpha_j^{n_0} := \sum_{l=0}^{k_j-1} \frac{\Delta t_{n_0+l}}{\Delta x}.$$

**Remark.** Note that a rarefaction wave  $\phi(x/t)$  solution of the scalar equation is such that  $f'(\phi(x/t)) - x/t = 0$  (continuous version of  $f'(u_j^{n_0}) \Delta t / \Delta x = 1$ ). This corresponds to  $\sum_{l=1}^{k_j-1} \frac{\Delta t_{n_0+l}}{\Delta x} \lambda_j^{n_0} = 1$  that is  $c_j^{n_0+k_j} = 0$ .

The following lemma will be useful to prove the convergence.

**Lemma 2.2.** Suppose that the exact solution is a rarefaction wave. Then for all  $j \in \mathbb{Z}$  and for all  $n \in \mathbb{N}$ , there exists  $n' > n$ , such that

$$|v(x_j, t_{n'}) - u_j^{n'}| = |v(x_j, 0) - u_j^0| + \mathcal{O}(\Delta t^2),$$

with

$$|v(x_j, t_k) - u_j^k| = |v(x_j, 0) - u_j^0| + \mathcal{O}(\Delta t), \quad n \leq k < n'.$$

**Proof.** By Taylor expansion at time  $t_{n_0+k_j}$  we have

$$v(j\Delta x, t_{n_0+k_j}) = v(j\Delta x, t_{n_0}) + \sum_{l=0}^{k_j-1} \Delta t_{n_0+l} v_t(j\Delta x, t_{n_0}) + \mathcal{O}\left(\left(\sum_{l=0}^{k_j-1} \Delta t_{n_0+l}\right)^2\right).$$

Using that  $v_t = -f(v)_x = -v_x f'(v)$  we get

$$v(j\Delta x, t_{n_0+k_j}) = v(j\Delta x, t_{n_0}) - \sum_{l=0}^{k_j-1} \Delta t_{n_0+l} v_x(j\Delta x, t_{n_0}) f'(v(j\Delta x, t_{n_0})) + \mathcal{O}\left(\left(\sum_{l=0}^{k_j-1} \Delta t_{n_0+l}\right)^2\right).$$

We can rewrite the previous equation:

$$\begin{aligned}
v(j\Delta x, t_{n_0+k_j}) &= v(j\Delta x, t_{n_0}) - \sum_{l=0}^{k_j-1} \Delta t_{n_0+l} \frac{v(j\Delta x, t_{n_0}) - v((j-1)\Delta x, t_{n_0})}{\Delta x} f'(v(j\Delta x, t_{n_0})) \\
&\quad + \mathcal{O}\left(\left(\sum_{l=0}^{k_j-1} \Delta t_{n_0+l}\right)^2\right) + \mathcal{O}\left(\Delta x \sum_{l=0}^{k_j-1} \Delta t_{n_0+l}\right).
\end{aligned}$$

If we denote by  $\phi(x/t)$  the exact self-similar continuous rarefaction wave, we formally have

$$v(j\Delta x, t_{n_0+k_j}) = \phi\left(\frac{j\Delta x}{t_{n_0+k_j}}\right)$$

and then

$$\begin{aligned}
\phi\left(\frac{j\Delta x}{t_{n_0+k_j}}\right) &= \phi\left(\frac{j\Delta x}{t_{n_0}}\right) - \sum_{l=0}^{k_j-1} \frac{\Delta t_{n_0+l}}{\Delta x} f'\left(\phi\left(\frac{j\Delta x}{t_{n_0}}\right)\right) \left[\phi\left(\frac{j\Delta x}{t_{n_0}}\right) - \phi\left(\frac{(j-1)\Delta x}{t_{n_0}}\right)\right] \\
&\quad + \mathcal{O}\left(\left(\sum_{l=0}^{k_j-1} \Delta t_{n_0+l}\right)^2\right) + \mathcal{O}\left(\Delta x \sum_{l=0}^{k_j-1} \Delta t_{n_0+l}\right).
\end{aligned}$$

When  $\sum_{l=0}^{k_j-1} \frac{\Delta t_{n_0+l}}{\Delta x} f'(\phi(j\Delta x/t_{n_0})) = 1$  (continuous version of (13)), we then we recover at order 2 (14):

$$|v(x_j, t_{n_0+k_j}) - u_j^{n_0+k_j}| = |v(x_j, t_{n_0}) - u_j^{n_0}| + \mathcal{O}(\Delta t^2).$$

That means that for the cell  $j$ , there exists a time such that the error in this cell is reduced to second order. This result has been proved for a particular time  $t_{n_0}$  and can be obviously extended to any positive time, by induction provided that no wave interaction occurs. This concludes the proof of the lemma.  $\square$

Again for all  $j \in \mathbb{Z}$  and for all time  $(t_{n_0:j_i})_{i \in \mathbb{N}}$  such that  $c_j^{n_0:j_i} = 0$ , the numerical solution satisfies

$$|v(x_j, t_{n_0:j_i}) - u_j^{n_0:j_i}| = |v(x_j, 0) - u_j^0| + \mathcal{O}(\Delta t^2).$$

However, between  $t_{n_0:j_i}$  and  $t_{n_0:j_{i+1}}$  because of (4), we have the following estimate:

$$|v(x_j, t_k) - u_j^k| = |v(x_j, 0) - u_j^0| + \mathcal{O}(\Delta t), \quad \forall k \in \{n_0:j_i + 1, \dots, n_0:j_{i+1} - 1\}.$$

In other words, in the cell  $j$ , an error in  $\Delta t^2$  is created not at each iteration, but at each  $n_0:j_k - n_0:j_{k-1}$  iterations. More generally for each  $j$  there exists  $0 < K_j = K_j(n) < n$  such that

$$|v(x_j, t_n) - u_j^n| = \sum_{k=0}^{K_j-1} |v(x_j, t_{n_0:j_k}) - u_j^{n_0:j_k}| + (\mathcal{O}(K_j \Delta t^2))$$

with  $t_{n_0:j_{K_j-1}} = t_n$ . The key point there is that  $K_j$  is strictly (and can be much) less than  $n$ , leading to a better approximation than classical order 1 schemes. Denoting  $K(n) = \sup_j K_j(n)$ , we then have by summing over  $j$  the existence two real constants  $c_2(t_n)$  and  $c_3 = c_3(TV(u^0))$

$$\|u^n - v(\cdot, t_n)\|_{L^1} \leq \|u^0 - v(\cdot, 0)\|_{L^1} + c_2(t_n)TV(u^0)\Delta t + c_3(TV(u^0))K(n)\Delta t^2, \quad (15)$$

where  $c(t_n)$  is bounded by a constant  $\bar{c}$ , for all  $n$ .

### Step 3. Conclusion.

As the solution of a hyperbolic equation of conservation law is a combination of shock and rarefaction waves and as we consider non-interacting waves, we can deduce that there exist  $c(t_n) > 0$  and  $d = d(TV(u^0)) > 0$  such that:

$$\|u^n - v(\cdot, t_n)\|_{L^1} \leq \|u^0 - v(\cdot, 0)\|_{L^1} + c(t_n)TV(u^0)\Delta t + dK(n)\Delta t^2.$$

That is, for step-like initial data the reservoir scheme is convergent.  $\square$

We now prove that for scalar equations the combination of flux schemes and reservoirs is total variation diminishing (TVD).



**Theorem 2.3.** *The reservoir technique is TVD for scalar hyperbolic equations of conservation laws (5) with non-interacting wave solutions and a strictly convex  $C^2$  flux,  $f$ :*

$$TV(u^{n+1}) = \sum_{j \in \mathbb{Z}} |u_j^{n+1} - u_{j-1}^{n+1}| \leq \sum_{j \in \mathbb{Z}} |u_j^n - u_{j-1}^n| = TV(u^n), \quad \forall n \in \mathbb{N}.$$

**Proof.** Recall first that under CFL condition, flux schemes are TVD [10]. If we denote by  $w^n$  a flux scheme solution and we suppose that at each iteration the time step is chosen such that the CFL number is equal to 1, we then have:

$$TV(w^{n+1}) \leq TV(w^n), \quad \forall n \in \mathbb{N}.$$

Now denoting by  $(n_{0;j})_j$  the indices such that  $c_j^{n_{0;j}} = 0$ , the reservoir scheme is defined in the cell  $j$  and for  $n$  greater than  $n_{0;j}$ , by:

$$u_j^{n+1} = \begin{cases} u_j^{n_{0;j}}, & \text{if } c_j^{n+1} > 0, \\ u_j^{n_{0;j}} - \sum_{l=0}^{n-n_{0;j}} \frac{\Delta t_{n_{0;j}+l}}{\Delta x} (g(u_j^{n_{0;j}}, u_{j+1}^{n+l}) - g(u_{j-1}^{n+l}, u_j^{n_{0;j}})), & \text{if } c_j^{n+1} = 0. \end{cases}$$

Then if we sum over  $j$

$$\sum_{j \in \mathbb{Z}} |u_j^{n+1} - u_{j-1}^{n+1}| = \sum_{j \in \mathbb{Z}} \left| u_j^{n_{0;j}} - \sum_{l=0}^{n-n_{0;j}} \frac{\Delta t_{n_{0;j}+l}}{\Delta x} (\cdots) \mathbf{1}_{c_j^{n+1}=0} - u_{j-1}^{n_{0;j-1}} + \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (\cdots) \mathbf{1}_{c_{j-1}^{n+1}=0} \right|.$$

Again, as we do not consider wave interactions only three situations can occur:

*Case 1. Shock waves.*

We have seen that the solution is updated only when local CFL numbers reach 1; this corresponds to an exact shock propagation. Then the total variation is then not modified in this case.

*Case 2. Constant states.*

Trivially the total variation is zero.

*Case 3. Rarefaction waves.*

This situation is more complicated. We will prove for right-rarefaction waves (by symmetry the result will be valid for left rarefaction waves), that the total variation diminishing using an inductive argument and the reservoir and counter features.

From time  $t = 0$  to  $t_1$  we have:

$$u_j^1 = \begin{cases} u_j^0, & \text{if } c_j^1 > 0, \\ u_j^0 - \frac{\Delta t_0}{\Delta x} (f(u_j^0) - f(u_{j-1}^0)), & \text{if } c_j^1 = 0, \end{cases}$$

and

$$u_{j-1}^1 = \begin{cases} u_{j-1}^0, & \text{if } c_{j-1}^1 > 0, \\ u_{j-1}^0 - \frac{\Delta t_0}{\Delta x} (f(u_{j-1}^0) - f(u_{j-2}^0)), & \text{if } c_{j-1}^1 = 0. \end{cases}$$

Then, as  $f$  is increasing (recall that we consider right-rarefaction waves) and using the fact that the time step is chosen such that local CFL numbers are less or equal to 1, we check easily that for right-rarefaction waves, we have:

$$|u_j^1 - u_{j-1}^1| \leq |u_j^0 - u_{j-1}^0|,$$

so that:

$$TV(u^1) = \sum_{j \in \mathbb{Z}} |u_j^1 - u_{j-1}^1| \leq \sum_{j \in \mathbb{Z}} |u_j^0 - u_{j-1}^0| = TV(u^0).$$

Suppose now that at time  $t_n$ :

$$TV(u^k) \leq TV(u^{k-1}), \quad 1 \leq k \leq n.$$

At time  $t_{n+1}$  we have the following equations:

$$\begin{cases} u_j^{n+1} = u_j^n - \sum_{l=0}^{n-n_{0;j}} \frac{\Delta t_{n_{0;j}+l}}{\Delta x} (f(u_j^{n_{0;j}}) - f(u_{j-1}^{n+l})) \mathbf{1}_{c_j^{n+1}=0}, \\ u_{j-1}^{n+1} = u_{j-1}^n - \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-2}^{n+l})) \mathbf{1}_{c_{j-1}^{n+1}=0}. \end{cases}$$

Then

$$\begin{aligned} |u_j^{n+1} - u_{j-1}^{n+1}| &= \left| u_j^n - u_{j-1}^n - \sum_{l=0}^{n-n_{0;j}} \frac{\Delta t_{n_{0;j}+l}}{\Delta x} (f(u_j^{n_{0;j}}) - f(u_{j-1}^{n_{0;j}+l})) \mathbf{1}_{c_j^{n+1}=0} \right. \\ &\quad \left. + \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-2}^{n_{0;j-1}+l})) \mathbf{1}_{c_{j-1}^{n+1}=0} \right|. \end{aligned}$$

Four situations can occur (corresponding to four kinds of cells):

- If  $c_j^{n+1} > 0$ ,  $c_{j-1}^{n+1} > 0$  then

$$u_j^{n+1} = u_j^n = u_j^{n_{0;j}}, \quad u_{j-1}^{n+1} = u_{j-1}^n = u_{j-1}^{n_{0;j-1}}.$$

In such cells  $j$ , we have:

$$|u_j^{n+1} - u_{j-1}^{n+1}| = |u_j^n - u_{j-1}^n|.$$

For such cells there is no increasing variation.

- If  $c_j^{n+1} = 0$ ,  $c_{j-1}^{n+1} > 0$  then

$$|u_j^{n+1} - u_{j-1}^{n+1}| = \left| u_j^n - u_{j-1}^n - \sum_{l=0}^{n-n_{0;j}} \frac{\Delta t_{n_{0;j}+l}}{\Delta x} (f(u_j^{n_{0;j}}) - f(u_{j-1}^{n_{0;j}+l})) \right|.$$

Then as  $f$  is increasing (right-rarefaction waves are considered) and by entropy condition the term inside the sum is positive and then:

$$- \sum_{l=0}^{n-n_{0;j}} \frac{\Delta t_{n_{0;j}+l}}{\Delta x} (f(u_j^{n_{0;j}}) - f(u_{j-1}^{n_{0;j}+l})) \leq 0.$$

We have (under CFL condition):

$$|u_j^{n+1} - u_{j-1}^{n+1}| \leq |u_j^n - u_{j-1}^n|.$$

- If now  $c_j^{n+1} > 0$ ,  $c_{j-1}^{n+1} = 0$ , then

$$|u_j^{n+1} - u_{j-1}^{n+1}| = \left| u_j^n - u_{j-1}^n + \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-2}^{n_{0;j-1}+l})) \right|.$$

And again the term inside the sum is positive, so that:

$$|u_j^{n+1} - u_{j-1}^{n+1}| \geq |u_j^n - u_{j-1}^n|.$$

At this stage, we cannot conclude anything about the total variation. For the  $(j-2)$  cell, we have:

$$u_{j-2}^{n+1} = u_{j-2}^n - \sum_{l=0}^{n-n_{0;j-2}} \frac{\Delta t_{n_{0;j-2}+l}}{\Delta x} (f(u_{j-2}^{n_{0;j-2}}) - f(u_{j-3}^{n_{0;j-2}+l})) \mathbf{1}_{c_{j-2}^{n+1}=0}.$$

Now let us denote by  $k$ , with  $k \geq 3$  the smallest index such that  $c_{j-k}^{n+1} > 0$  with  $c_{j-k'}^{n+1} = 0$  for  $k' \in \{1, \dots, k-1\}$ . To simplify the notations let us suppose that  $k = 3$ . Then if  $c_{j-2}^{n+1} > 0$ :

$$\begin{aligned} |u_{j-1}^{n+1} - u_{j-2}^{n+1}| &= \left| u_{j-1}^n - u_{j-2}^n - \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-2}^{n_{0;j-1}+l})) \right| \\ &\leq |u_{j-1}^n - u_{j-2}^n|. \end{aligned}$$

Defining  $A$  by:

$$A = u_{j-1}^n - u_{j-2}^n - \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-2}^{n_{0;j-1}+l})),$$

we obtain by Taylor expansion (rarefaction waves are considered here):

$$\begin{aligned}
A &= u_{j-1}^n - u_{j-2}^n - \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-1}^{n_{0;j-1}+l})) \\
&\quad - f'(u_{j-1}^{n_{0;j-1}+l})(u_{j-2}^{n_{0;j-1}+l} - u_{j-1}^{n_{0;j-1}+l}) - f''(u_{j-1}^{n_{0;j-1}+l})(u_{j-2}^{n_{0;j-1}+l} - u_{j-1}^{n_{0;j-1}+l})^2/2 \\
&\quad + \mathcal{O}((u_{j-2}^{n_{0;j-1}+l} - u_{j-1}^{n_{0;j-1}+l})^3).
\end{aligned}$$

Then using the fact that

$$\sum_{l=0}^{n-n_{0;j-1}} \lambda_{j-1}^{n_{0;j-1}+l} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} = 1,$$

we deduce that  $A$  is positive

$$A = \frac{1}{2} f''(u_{j-1}^{n_{0;j-1}+l})(u_{j-2}^{n_{0;j-1}+l} - u_{j-1}^{n_{0;j-1}+l})^2 + \mathcal{O}((u_{j-2}^{n_{0;j-1}+l} - u_{j-1}^{n_{0;j-1}+l})^3) \geq 0.$$

Then we can deduce that:

$$|u_{j-1}^{n+1} - u_{j-2}^{n+1}| + |u_j^{n+1} - u_{j-1}^{n+1}| \leq |u_{j-1}^n - u_{j-2}^n| + |u_j^n - u_{j-1}^n|,$$

as:

$$\begin{aligned}
u_j^{n+1} - u_{j-1}^{n+1} &= u_j^n - u_{j-1}^n + \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-2}^{n_{0;j-1}+l})) > 0, \\
u_{j-1}^{n+1} - u_{j-2}^{n+1} &= u_{j-1}^n - u_{j-2}^n - \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-2}^{n_{0;j-1}+l})) > 0.
\end{aligned}$$

In the situation  $k > 3$  with  $c_{j-k'}^{n+1} > 0$  with  $c_{j-k}^{n+1} = 0$  for  $k' \in \{1, \dots, k-1\}$  we would have:

$$\sum_{l=0}^{k-2} |u_{j-l}^{n+1} - u_{j-l-1}^{n+1}| \leq \sum_{l=0}^{k-2} |u_{j-l}^n - u_{j-l-1}^n|.$$

- The last case corresponds to  $c_j^{n+1} = 0$ ,  $c_{j-1}^{n+1} = 0$ . That is:

$$\begin{cases}
u_j^{n+1} = u_j^n - \sum_{l=0}^{n-n_{0;j}} \frac{\Delta t_{n_{0;j}+l}}{\Delta x} (f(u_j^{n_{0;j}}) - f(u_{j-1}^{n_{0;j}+l})), \\
u_{j-1}^{n+1} = u_{j-1}^n - \sum_{l=0}^{n-n_{0;j-1}} \frac{\Delta t_{n_{0;j-1}+l}}{\Delta x} (f(u_{j-1}^{n_{0;j-1}}) - f(u_{j-2}^{n_{0;j-1}+l})).
\end{cases}$$

But as  $f'$  is strictly positive that necessarily involves that  $n_{0;j-1} < n_{0;j}$  and the following situation can occur:

$$|u_j^{n+1} - u_{j-1}^{n+1}| \geq |u_j^n - u_{j-1}^n|.$$

But as above (third points) there will also exist  $k$  such that:

$$\sum_{l=0}^k |u_{j-l}^{n+1} - u_{j-l-1}^{n+1}| \leq \sum_{l=0}^k |u_{j-l}^n - u_{j-l-1}^n|.$$

Finally for rarefaction waves we globally have:

$$TV(u^{n+1}) = \sum_{j \in \mathbb{Z}} |u_j^{n+1} - u_{j-1}^{n+1}| \leq \sum_{j \in \mathbb{Z}} |u_j^n - u_{j-1}^n| = TV(u^n). \quad \square$$

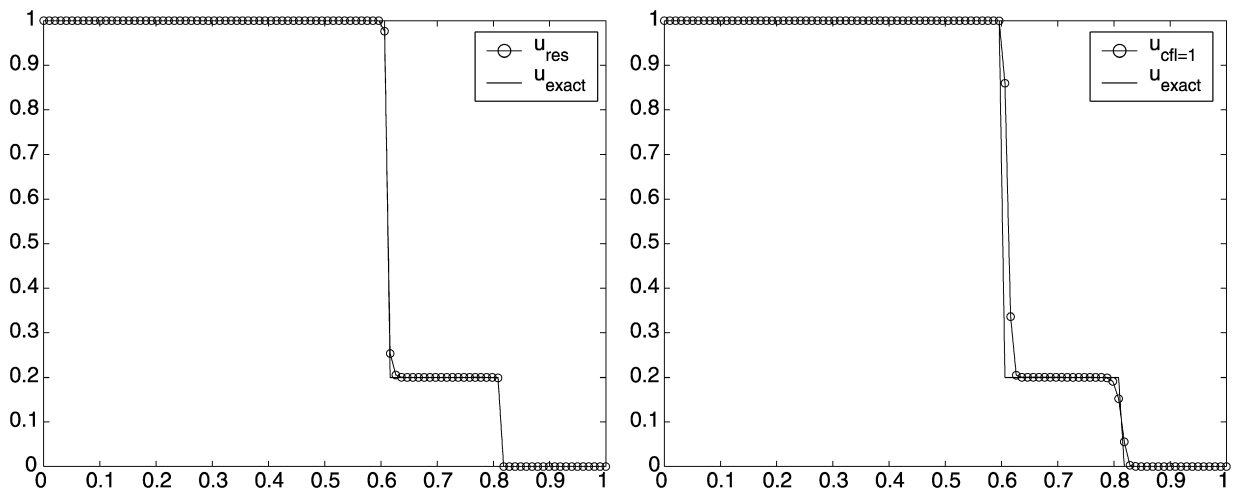


Fig. 1. Numerical solution at time  $T = 0.6$ : Reservoir and upwind at “CFL = 1”.

### 2.2.1. On wave interactions

For the sake of simplicity, we have considered in the above studies, non-interacting wave solutions, although wave interaction is fundamental, especially for long time simulations. This has allowed us to focus on the main specificity of the reservoir technique: very low diffusion of propagating shock and rarefaction waves. Moreover as it is well known, an accurate treatment of interactions is mainly the fact of a good choice of the (linear or nonlinear) Riemann solver. The same manner the Colella–Glaz Riemann solver [3] allowed us to treat very accurately wave decompositions for hyperbolic systems.

In order to be exhaustive 3 kinds of interaction should be considered: shock-shock, shock-rarefaction and rarefaction-rarefaction interactions. We can in fact simply be deduced from the above study a corollary of Theorem 2.2:

**Corollary 2.1.** Suppose that  $i_1$  shock-shock interactions,  $i_2$  shock-rarefaction interactions and  $i_3$  rarefaction-rarefaction interactions have occurred for  $0 \leq t \leq t_n$ . Then there exist two positive constants  $c(t_n)$  and  $d = d(TV(u^0))$  such that:

$$\|u^n - v(\cdot, t_n)\|_{L^1} \leq \|u^0 - v(\cdot, 0)\|_{L^1} + c(t_n)(i_1 + i_2 + i_3)TV(u^0)\Delta t + d(i_2 + i_3)K(n)\Delta t^2 \quad (16)$$

and there exists  $\bar{c} > 0$  such that  $\sup_{n \in \mathbb{N}} c(t_n) \leq \bar{c}$  and  $K(n) < n$ .

**Principle of the proof.** Let us consider first the shock-shock interaction case: two situations are possible. If two shock waves interact at time  $t^* \leq t_n$  such that their respective reservoir and counter is zero (that is interaction of non-diffused shock waves), the new created shock will propagate following the same process as described above (with diffusion production then cancellation), and (12) will still be valid. Now, if at least one of the two shock waves is diffused at the interaction time  $t^*$ , its corresponding reservoir will have not been emptied before the interaction and some numerical diffusion of order  $\mathcal{O}(\Delta t)$  (11) (but independently of  $t^*$ ) will not be canceled after the creation of the new shock wave. In that case, the global error will increase during the interaction following (16).

Using (15), this principle can be extended the same manner to shock-rarefaction interactions, or rarefaction-rarefaction interactions, considering that the interaction could occur when both reservoirs are empty, only one, or both full. We clearly note that as expected, the interactions lead potentially to increase the numerical diffusion of the solution.  $\square$

### 2.2.2. Numerical tests

In order to illustrate Theorem 2.2 we propose some numerical tests on Burgers' equation:  $v_t + v v_x = 0$ ,  $v(x, 0) = v_0(x)$ .

- The first test consists of solving with the reservoir technique coupled with a classical upwind scheme at “CFL = 1” (CFL = 0.999 in practice), Burgers' equation in  $[0, 1]$  with the following initial data:

$$v_0(x) = \begin{cases} 1, & \text{if } x < 1/4, \\ 0.2, & \text{if } 1/4 \leq x < 3/4, \\ 0, & \text{if } x \geq 3/4. \end{cases}$$

The final time is  $T = 0.6$ , the space step is  $\Delta x = 0.01$ . The numerical results are obtained in Fig. 1. Then we present in Fig. 2 the  $L^1$ -error between the exact and reservoir solution:  $(t, \|u_{\text{res}}(\cdot, t) - u_{\text{exact}}(\cdot, t)\|_{L^1([0, 1])})$  for  $0 \leq t \leq T$ . As proven above, the reservoir technique is convergent for large times (in particular the numerical error is bounded by a constant times the space step), which is not true for classical finite volume schemes.

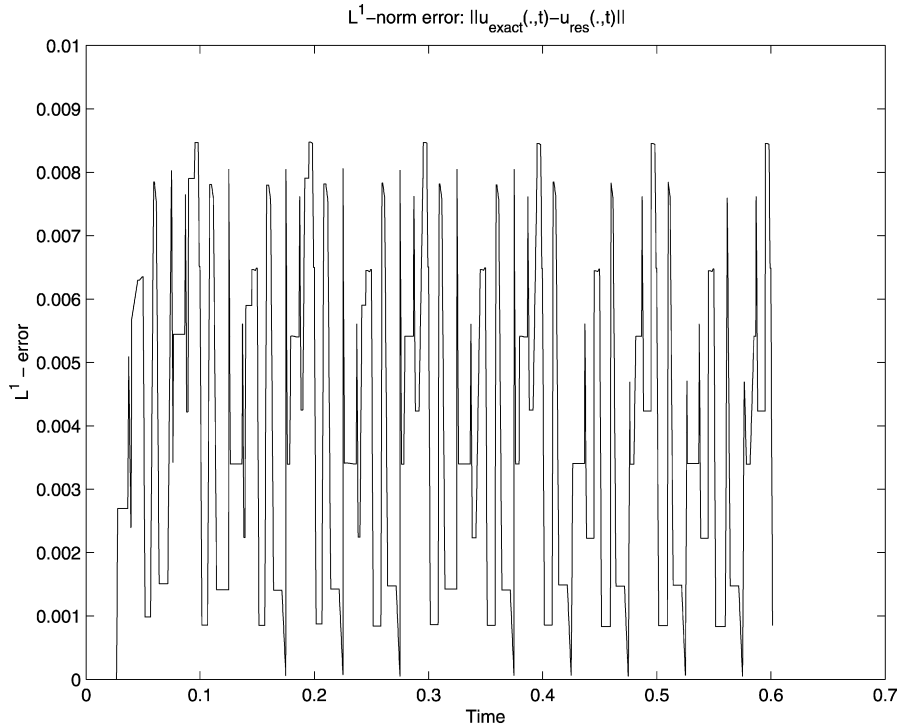


Fig. 2.  $L^1$ -norm error function of time.

- We now propose a numerical test involving two rarefaction waves. We compare the classical upwind scheme at “CFL = 1” (CFL = 0.999 precisely), with the reservoir technique. The domain is  $[0, 4]$  and we again solve Burgers’ equation. The initial data is given by:

$$v_0(x) = \begin{cases} 0, & \text{if } x < 0.4, \\ 0.8, & \text{if } 0.4 \leq x < 2, \\ 1, & \text{if } x \geq 2. \end{cases}$$

We represent in Fig. 3, the exact, reservoir and CFL = 1 solutions and in Fig. 4 the error in  $L^1$ -norm as a function of time, between the exact and the reservoir solutions and between the exact and the upwind solutions at “CFL = 1”, on a 200 cell mesh. The results are in accordance with Theorem 2.2. Indeed as expected the long time divergence of the reservoir technique is slower ( $K(n) < n$ , Theorem 2.2) than classical order one method.

### 2.3. Linear systems

In this section we study the convergence of the reservoir technique approximating linear systems of conservation laws without wave interaction.

#### 2.3.1. Reservoir scheme for linear systems

Consider the linear hyperbolic system with  $m$  equations:

$$\begin{cases} \partial_t V + \partial_x(AV) = 0, & x \in \mathbb{R}, t \geq 0, \\ V(x, 0) = V_0(x), & V_0 \in BV(\mathbb{R}), \end{cases} \quad (17)$$

with  $V = (v_1, \dots, v_m)$ ,  $A \in \mathcal{M}_m(\mathbb{R})$ , where  $\mathcal{M}_m(\mathbb{R})$  is the set of matrices, diagonalizable in  $\mathbb{R}$  with eigenvalues  $(\lambda_1, \lambda_2, \dots, \lambda_m)$ . To obtain a CFL equal to one on each characteristic field, a specific choice of time steps  $\Delta t_n$  is necessary. In this goal, we introduce  $m$  vectorial reservoirs  $R_{1;j}, \dots, R_{m;j} \in \mathbb{R}^m$  associated to the cell  $j$  and CFL counters  $c_{k;j} \in [0, 1]$ ,  $k = 1, \dots, m$ , initialized to zero. We denote by  $V_R^k(V, W)$  the solution of the Riemann problem with left state  $V$  and right state  $W$  which lies between the  $k$ th and the  $(k+1)$ th wave with  $V_R^0(V, W) := V$ , and  $V_R^m(V, W) := W$  by convention. For convenience we set in the following

$$c_{k;j}^{n+1} = c_{k;j}^n + |\lambda_k| \frac{\Delta t_n}{\Delta x}.$$

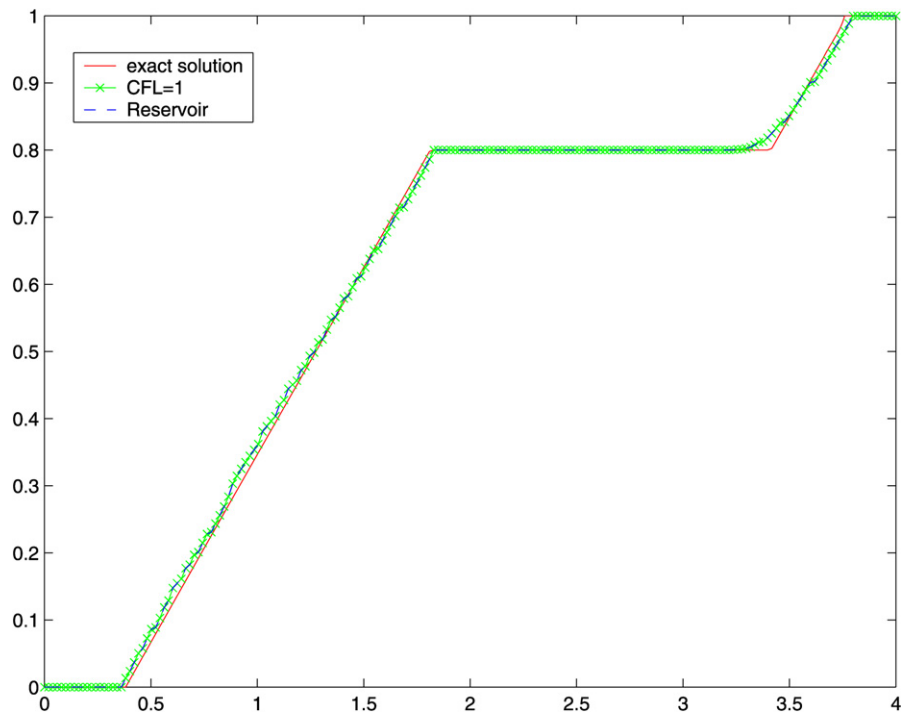


Fig. 3. Exact, reservoir and CFL = 1 solutions at time  $T = 1.8$ .

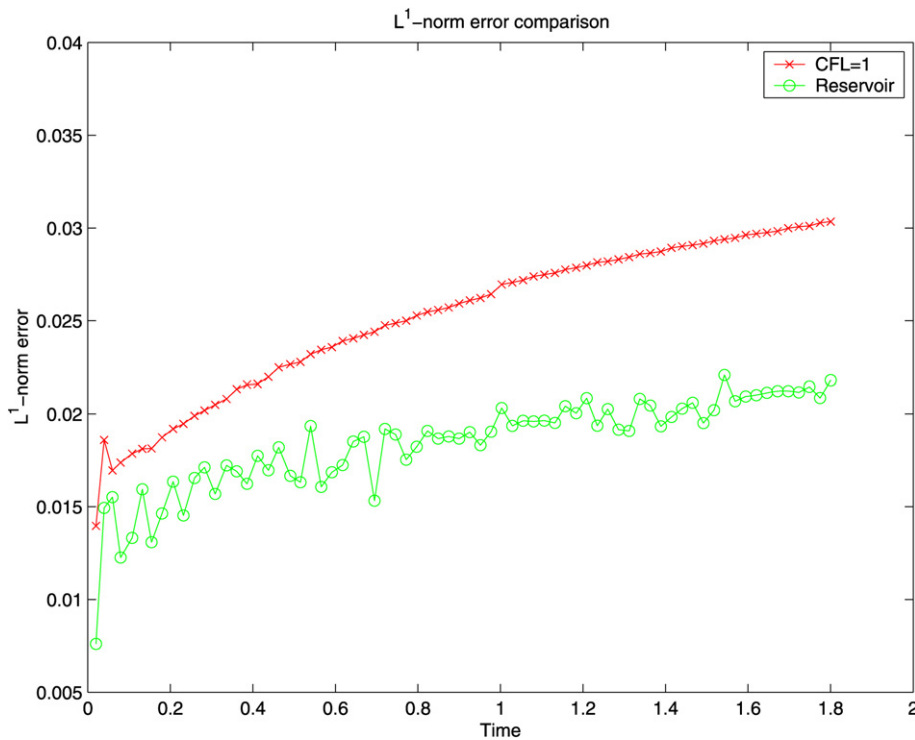


Fig. 4.  $L^1$ -norm error as a function of time.

At each time step we fill up the reservoirs  $R_{k,j}$  with the current numerical flux difference, upwinding regarding the sign of  $\lambda_k$ . More precisely, we take for  $\lambda_k < 0$

$$\begin{pmatrix} \tilde{V}_{k;j}^{n+1} \\ \tilde{C}_{k;j}^{n+1} \\ \tilde{R}_{k;j}^{n+1} \end{pmatrix} = \begin{cases} \begin{pmatrix} 0 \\ C_{k;j}^n + |\lambda_k| \frac{\Delta t_n}{\Delta x} \\ R_{k;j}^n - \frac{\Delta t_n}{\Delta x} (F(V_R^k(V_j^n, V_{j+1}^n)) - F(V_R^{k-1}(V_j^n, V_{j+1}^n))) \end{pmatrix}, & \text{if } C_{k;j}^{n+1} < 1, \\ \begin{pmatrix} 0 \\ 0 \\ R_{k;j}^n - \frac{\Delta t_n}{\Delta x} (F(V_R^k(V_j^n, V_{j+1}^n)) - F(V_R^{k-1}(V_j^n, V_{j+1}^n))) \end{pmatrix}, & \text{if } C_{k;j}^{n+1} = 1, \end{cases} \quad (18)$$

whereas, for  $\lambda_k > 0$ , we upwind on the right side

$$\begin{pmatrix} \tilde{V}_{k;j}^{n+1} \\ \tilde{C}_{k;j}^{n+1} \\ \tilde{R}_{k;j}^{n+1} \end{pmatrix} = \begin{cases} \begin{pmatrix} 0 \\ C_{k;j}^n + \lambda_k \frac{\Delta t_n}{\Delta x} \\ R_{k;j}^n - \frac{\Delta t_n}{\Delta x} (F(V_R^k(V_{j-1}^n, V_j^n)) - F(V_R^{k-1}(V_{j-1}^n, V_j^n))) \end{pmatrix}, & \text{if } C_{k;j}^{n+1} < 1, \\ \begin{pmatrix} 0 \\ 0 \\ R_{k;j}^n - \frac{\Delta t_n}{\Delta x} (F(V_R^k(V_{j-1}^n, V_j^n)) - F(V_R^{k-1}(V_{j-1}^n, V_j^n))) \end{pmatrix}, & \text{if } C_{k;j}^{n+1} = 1. \end{cases} \quad (19)$$

We update the solution by taking

$$V_j^{n+1} = V_j^n + \sum_{k=1}^m \tilde{V}_{k;j}^{n+1}.$$

As before, the time step  $\Delta t_n$  must be chosen according to the classical stability condition and as big as possible. This leads to the natural choice

$$\Delta t_n = \min_{j,k} \left( (1 - c_{k;j}^n) \frac{\Delta x}{|\lambda_k|} \right). \quad (20)$$

### 2.3.2. Numerical analysis

We consider here

$$V_t + AV_x = 0, \quad V(x, 0) = V_0(x), \quad V_0 \in L^1(\mathbb{R}) \cap L^\infty(\mathbb{R}) \cap BV(\mathbb{R}), \quad (21)$$

with  $A$  being a diagonalizable matrix such that  $\lambda_1 < \lambda_2 < \dots < \lambda_m$ . The initial data is given by:

$$V_j^0 = \begin{cases} V_l := \sum_{i=1}^m \alpha_i \Gamma_i, & \text{if } j \leq j_0, \\ V_r := \sum_{i=1}^m \beta_i \Gamma_i, & \text{if } j > j_0. \end{cases} \quad (22)$$

Recall that an explicit flux scheme approximating a linear system of conservation law (21), is convergent (in particular stable) under a CFL condition  $\max_{i \in \{1, \dots, m\}} |\lambda_i| \Delta t / \Delta x \leq 1$ . Again we will use the fact that the reservoir technique combined with a flux scheme is equivalent to choosing particular time steps.

**Theorem 2.4.** *A flux scheme combined with the reservoir technique, approximating a linear hyperbolic system of conservation laws (21)–(22) is convergent, and there exists  $c > 0$  such that, at all time  $t_n \in \mathbb{R}_+$ :*

$$\|V(\cdot, t_n) - V^n\|_{L^1} \leq \|V(\cdot, 0) - V^0\|_{L^1} + c \Delta t.$$

Moreover for rational eigenvalues and for all time  $t_n$ , there exists a time  $t_{n'} \geq t_n$  such that the solution is exact at the discrete level, that is:

$$\|V(\cdot, t_{n'}) - V^{n'}\|_{L^1} = \|V(\cdot, 0) - V^0\|_{L^1}.$$

The numerical solution at time  $t_n$  in the cell  $j$  that is  $V_j^n$  is also denoted by  $(V_{j;l}^n)_{(1, \dots, m)}$  where  $V_{j;l}^n$  denotes the  $l$ th component of the solution in cell  $j$  at time  $t_n$ .

**Proof.** We first state an important lemma:

**Lemma 2.3.** *For every  $j \in \mathbb{Z}$ ,  $\forall l \in \{1, \dots, m\}$ ,  $\forall n > 0$ ,  $\exists n' > n$ ,*

$$\sum_j |V_l(x_j, t_{n'}) - V_{j;l}^{n'}| = \sum_j |V_l(x_j, 0) - V_{j;l}^0|,$$

with

$$\sum_j |V_l(x_j, t_k) - V_{j,l}^k| = \sum_j |V_l(x_j, 0) - V_{j,l}^0| + \mathcal{O}(\Delta t), \quad n \leq k < n'.$$

Namely, we have for all  $j \in \mathbb{Z}$ , for all  $l \in \{1, \dots, m\}$ , and for all  $n_{0,j}$  such that  $c_j^{n_{0,j}} = 0$ , there exists  $k_j > 0$ , such that  $c_j^{n_{0,j}+k_j} = 0$  and

$$\sum_j |V_l(x_j, t_{n_{0,j}+k_j}) - V_{j,l}^{n_{0,j}+k_j}| = \sum_j |V_l(x_j, 0) - V_{j,l}^0|,$$

with

$$\sum_j |V_l(x_j, t_{n_{0,j}+k}) - V_{j,l}^{n_{0,j}+k}| = \sum_j |V_l(x_j, 0) - V_{j,l}^0| + \mathcal{O}(\Delta t), \quad n \leq k < k_j.$$

**Proof.** We suppose that  $\lambda_i \geq 0$ , for all  $1 \leq i \leq m$  and that  $\lambda_m/\lambda_1 \leq 2$ . These assumptions do not remove any theoretical difficulty, but allow us to simplify the notations. In particular the second one ensures us that for all  $N$  in  $\mathbb{N}$ , from the time iterations  $N \times m$  to  $(N+1) \times m$  each counter (recall that there are  $m$  counters) is updated one and only one time, what then simplifies greatly the indexation.

Suppose that the current time step index is equal to  $n$  with  $c_{k,j}^n = 0$  for all  $k$  in  $\{1, \dots, m\}$  and all  $j$  in  $\mathbb{Z}$  (such an  $n$  exists, for instance  $n = 0$ ). We consider an initial data given by (22). Then at each time step we update the solution in each cell in the following way:

$$V_{j+1}^{n+1} = V_{j+1}^n - \frac{\Delta t_n}{\Delta x} A(V_R^{m+1}(V_j^n, V_{j+1}^n) - V_R^m(V_j^n, V_{j+1}^n)) = V_{j+1}^n - \frac{\Delta t_n}{\Delta x} A(V_{j+1}^n - V_R^m(V_j^n, V_{j+1}^n)).$$

Without loss of generality, we can assume that  $A = \text{diag}(\lambda_1, \dots, \lambda_m)$ . First,  $\Delta t_n$  is chosen such that  $c_{m,j+1}^{n+1} = \lambda_m \Delta t_n / \Delta x = 1$  that is  $c_{m,j+1}^{n+1} = 0$  and as

$$V_{j+1}^n - V_R^m(V_j^n, V_{j+1}^n) = (\beta_m - \alpha_m) \mathbf{r}_m,$$

then

$$V_{j+1}^{n+1} = V_{j+1}^n - (\beta_m - \alpha_m) \mathbf{r}_m = \sum_{i=1}^{m-1} \beta_i \mathbf{r}_i + \alpha_m \mathbf{r}_m.$$

This corresponds to an exact propagation at the discrete level for the  $m$ th wave, that is without numerical diffusion.

Now, for the  $(m-p)$ th characteristic wave, with  $0 \leq p \leq m-1$ :

$$V_{j+1}^{n+p+1} = V_{j+1}^{n+p} - \sum_{i=0}^p \frac{\Delta t_{n+i}}{\Delta x} A(V_R^{m-p+1}(V_j^{n+p}, V_{j+1}^{n+p}) - V_R^{m-p}(V_j^{n+p}, V_{j+1}^{n+p})),$$

with

$$V_R^{m-p+1}(V_j^{n+p}, V_{j+1}^{n+p}) - V_R^{m-p}(V_j^{n+p}, V_{j+1}^{n+p}) = (\beta_{m-p} - \alpha_{m-p}) \mathbf{r}_{m-p}.$$

As  $c_{m-p,j+1}^{n+p} = \lambda_{m-p} \sum_{i=0}^p \Delta t_{n+i} / \Delta x = 1$ :

$$V_{j+1}^{n+p+1} = V_{j+1}^{n+p} - \sum_{i=0}^p \frac{\Delta t_{n+i}}{\Delta x} \lambda_{m-p} (\beta_{m-p} - \alpha_{m-p}) \mathbf{r}_{m-p} = \sum_{i=p}^m \alpha_i \mathbf{r}_i + \sum_{i=1}^{p-1} \beta_i \mathbf{r}_i.$$

This proves that for each component  $p \leq m$  the propagation of the  $p$ th wave is exact. Finally by induction, we can summarize the process by:

$$V_{j+1}^{n+m} = V_{j+1}^n - \sum_{i=0}^{m-1} \frac{\Delta t_{n+i}}{\Delta x} \lambda_{m-i} (\beta_{m-i} - \alpha_{m-i}) \mathbf{r}_{m-i} = V_j^n.$$

This process can then easily be extended for all  $t$ .  $\square$

This lemma guarantees us that at time  $t$ , for each characteristic wave, there exists a time  $t' \geq t$  such that the propagation is exact. However this does not allow us to conclude that for all  $t$  there exists a time  $t' \geq t$  such that the discrete solution is exact (without numerical diffusion). That is that the propagation is exact for all characteristic waves all together. First by simple extension of the previous lemma we have



**Remark.** If for all  $i \in \{1, \dots, m-1\}$   $\lambda_{i+1} = 2^i \lambda_1$ , then the discrete solution of the reservoir scheme is exact every  $2^m$  iterations. More precisely the propagation of the  $l$ th characteristic wave with  $l \in \{1, \dots, m\}$ , of the discrete solution is exact every  $2^l$  iterations.

To generalize this idea let us establish the following lemma:

**Lemma 2.4.** Suppose that  $\lambda_i \in \mathbb{Z}^*$ , for all  $i \in \{1, \dots, m\}$  and denote by  $P_{1,\dots,m}$  the following least common multiple:

$$P_{1,\dots,m} = \ell cm(|\lambda_1|, \dots, |\lambda_m|).$$

Then the discrete solution is exact every  $P_{1,\dots,m}$  iterations.

**Proof.** Let us denote by  $t_n$  the current time, such that all counters are zero (for example  $n = 0$ ). We then define  $(\rho_i)_{i \in \{1, \dots, m-1\}}$ :

$$\rho_i = \left\lfloor \frac{\lambda_{i+1}}{\lambda_i} \right\rfloor \in \mathbb{Q}^*, \quad \forall i \in \{1, \dots, m\},$$

and  $\delta t_i$  (CFL = 1 for the component  $i$ ):

$$\delta t_i = \frac{\Delta x}{|\lambda_i|}, \quad \forall i \in \{1, \dots, m\}.$$

Note that  $\delta t_i = \rho_i \delta t_{i+1}$  and recall that at times  $t_n + N \delta t_i$ ,  $N \in \mathbb{N}$ , the propagation of the  $i$ th wave is exact.

As  $\lambda_m$  is the fastest velocity  $\Delta t_{n+1} = \delta t_m = \Delta x / |\lambda_m|$ .

- If  $\rho_{m-1} \in \mathbb{N}^*$ , at time iterations  $[n, n + \rho_{m-1}[$  the  $m$ th reservoir is the only one to be updated, with zero diffusion on the associated characteristic wave. Then by the definition of the scheme at time  $t_{n+\rho_{m-1}}$  the  $m$ th and  $(m-1)$ th reservoirs are updated and the diffusion for the corresponding characteristic waves is zero.
- If now  $\rho_{m-1} \in \mathbb{Q}^* - \mathbb{N}$  and denoting by  $P_{m,m-1} = \ell cm(|\lambda_m|, |\lambda_{m-1}|)$  then we can assert that at times  $t_n + N P_{m,m-1} \delta t_m$ ,  $N \in \mathbb{N}^*$  the propagation is exact for these two characteristic waves.

Using similar arguments, for all  $k$  and  $l$  with  $1 \leq k < l < m$ , the propagation of the discrete  $k$ th and  $l$ th characteristic waves is exact at times  $t_n + N P_{l,k} \delta t_l$  with  $P_{l,k} = \ell cm(|\lambda_k|, |\lambda_l|)$  and  $N \in \mathbb{N}^*$ . By induction on the wave indices, we easily deduce that at times  $t_n + N P_{1,\dots,m} \delta t_m$ ,  $N \in \mathbb{N}^*$  the propagation is exact for all characteristic waves, that is the solution is exact at the discrete level.  $\square$

In fact we can even easily extend this result to rational eigenvalues.

**Lemma 2.5.** Supposing that the eigenvalues are rational, then for all time iteration  $n$ , there exists  $n' > n$  such that the discrete solution is exact at time  $t_{n'}$ .

**Proof.** Let us write  $\lambda_i$  as

$$\lambda_i = p_i / q_i, \quad i \in \{1, \dots, m\}, \quad p_i, q_i \text{ relatively prime integers},$$

and denote by  $R_{1,\dots,m}$  the following least common multiple:

$$R_{1,\dots,m} = \ell cm(|p_1|, \dots, |p_m|).$$

With the same arguments as above we prove that the solution is exact at the discrete level at least every  $R_{1,\dots,m}$  iterations. For all  $n$ , it is sufficient to define  $n'$  as a multiple of  $R_{1,\dots,m}$  greater than  $n$ .  $\square$

We can then deduce from Lemma 2.3 that for all  $n \in \mathbb{N}$  there exists  $c > 0$  such that:

$$\|V(\cdot, t_n) - V^n\|_{L^1} \leq \|V(\cdot, 0) - V^0\|_{L^1} + c \Delta t.$$

And from Lemma 2.5 we can add that for rational eigenvalues, there exists  $\tilde{n}$  such that  $c_{k,j}^{\tilde{n}} = 1$  for all  $k$  in  $\{1, \dots, m\}$  and:

$$\|V(\cdot, t_{\tilde{n}}) - V^{\tilde{n}}\|_{L^1} = \|V(\cdot, 0) - V^0\|_{L^1}. \quad \square$$

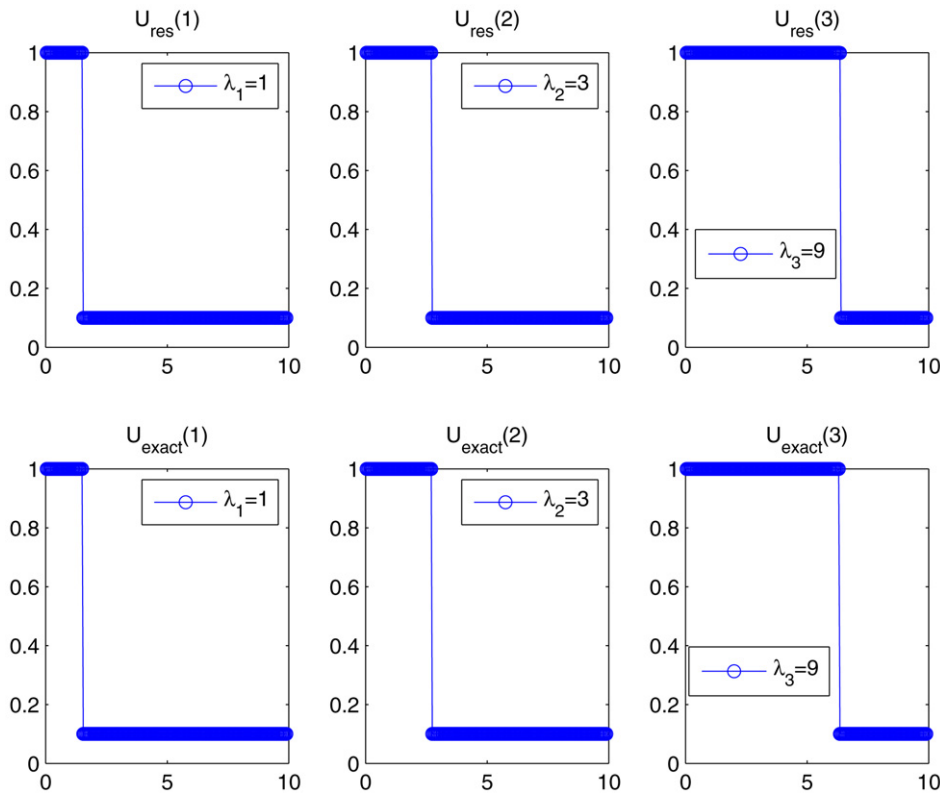


Fig. 5. Reservoir and exact solutions at time  $T = 0.6$  (rational eigenvalues).

### 2.3.3. Numerical tests

The previous results guarantee an exact propagation at the discrete level for rational eigenvalues. However we conjecture that the result can be extended to all real eigenvalues as it will be shown in the following numerical example. We solve the following system:

$$U_t + DU_x = 0,$$

where we suppose that  $D = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$  is a diagonal matrix (the system is then not coupled), but we solve it using one single time step (corresponding to a “CFL = 1” for the largest eigenvalue). The space step is given by  $\Delta x = 0.05$ , the domain is  $[0, 10]$ , the final time is  $T = 0.6$ , and the initial data is:

$$U_0 = \begin{cases} (1, 1, 1)^T, & \text{if } x \leq 1, \\ (0.1, 0.1, 0.1)^T, & \text{if } x > 1. \end{cases}$$

We consider two situations:

- First we suppose that the eigenvalues are rational, given respectively by  $\lambda_1 = 1$ ,  $\lambda_2 = 3$ ,  $\lambda_3 = 9$ . We represent in Figs. 5, 6 the reservoir and exact solution and the exact and upwind solution at “CFL = 1” (CFL = 0.999 in practice) that is  $\Delta t = \Delta x / \lambda_3$ . Fig. 7 gives the  $L^1$ -norm error between the exact and the reservoir solutions. As expected, the reservoir solution is convergent for large times but not the upwind solution as it is well known, Fig. 8.
- We now consider irrational eigenvalues, in order to prove that even in this situation the reservoir technique remains very accurate. The eigenvalues are given by  $\lambda_1 = \pi/10$ ,  $\lambda_2 = \pi$ ,  $\lambda_3 = 3\pi$ . Results are presented in Figs. 9 and 10, showing that the reservoir technique performs well (no diffusion) even when the eigenvalues are irrational. More generally our numerical tests have never made appear any difficulty regarding irrational eigenvalues.

## 3. Conclusion

In this paper, we have proven for a general class of initial data, the convergence and some important features of the reservoir technique for nonlinear hyperbolic equations and linear systems of conservation laws. For non-interacting shock wave solutions, we have in particular proven the *long time convergence* of the reservoir technique. In a forthcoming paper, the convergence of the reservoir technique for nonlinear hyperbolic systems of conservation law will be studied. For nonlinear

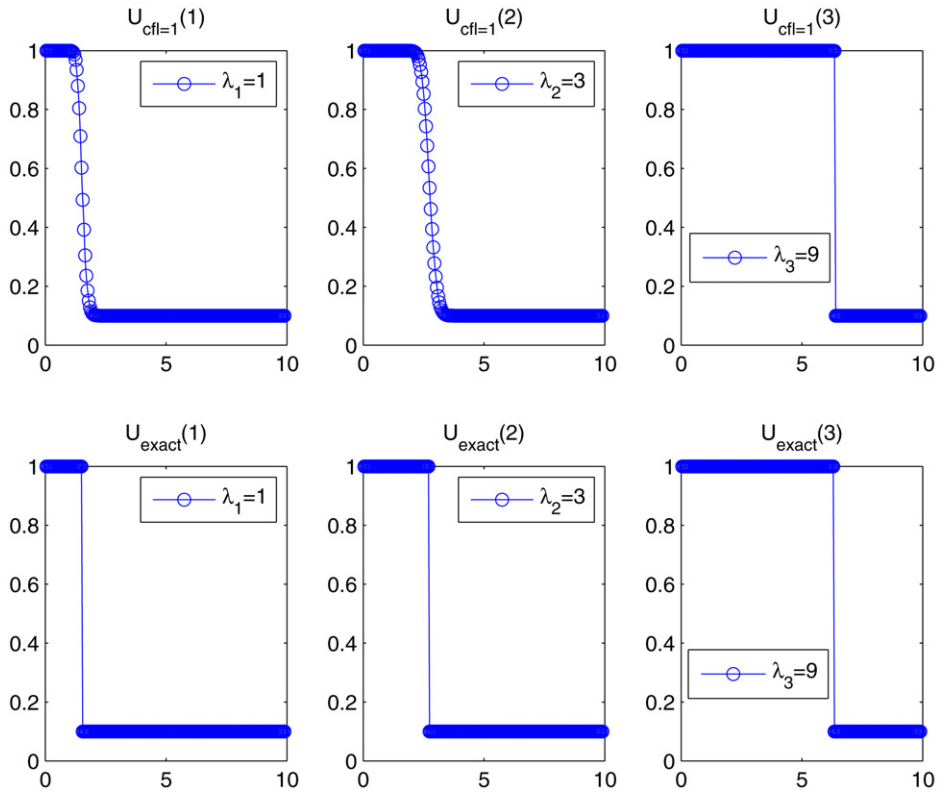


Fig. 6. Upwind at “CFL = 1” and exact solutions at time  $T = 0.6$  (rational eigenvalues).

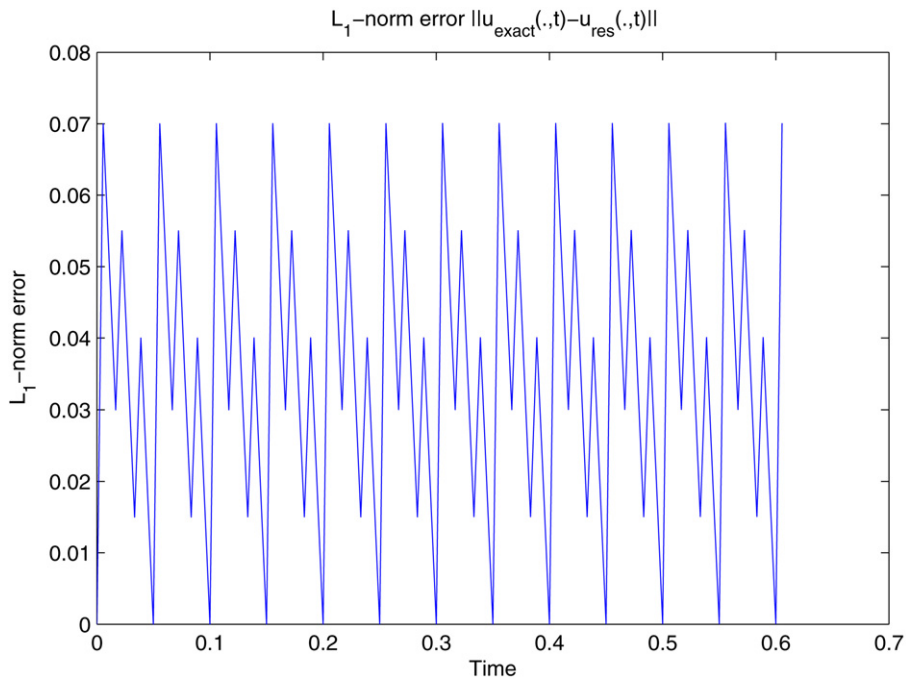


Fig. 7. Reservoir –  $L^1$ -norm error as a function of time (rational eigenvalues).

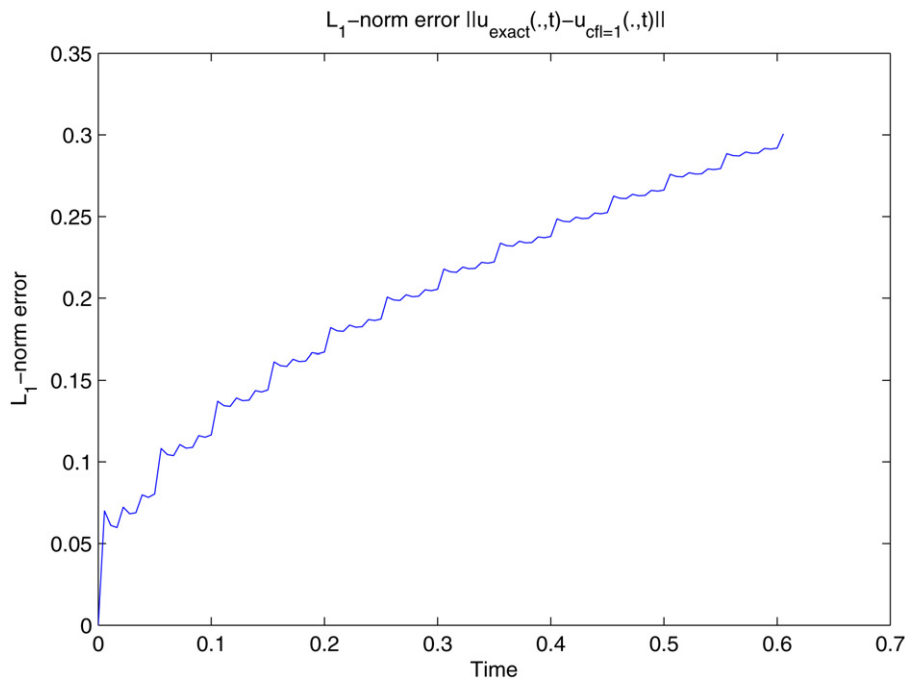


Fig. 8. Upwind at “CFL = 1” –  $L^1$ -norm error as a function of time.

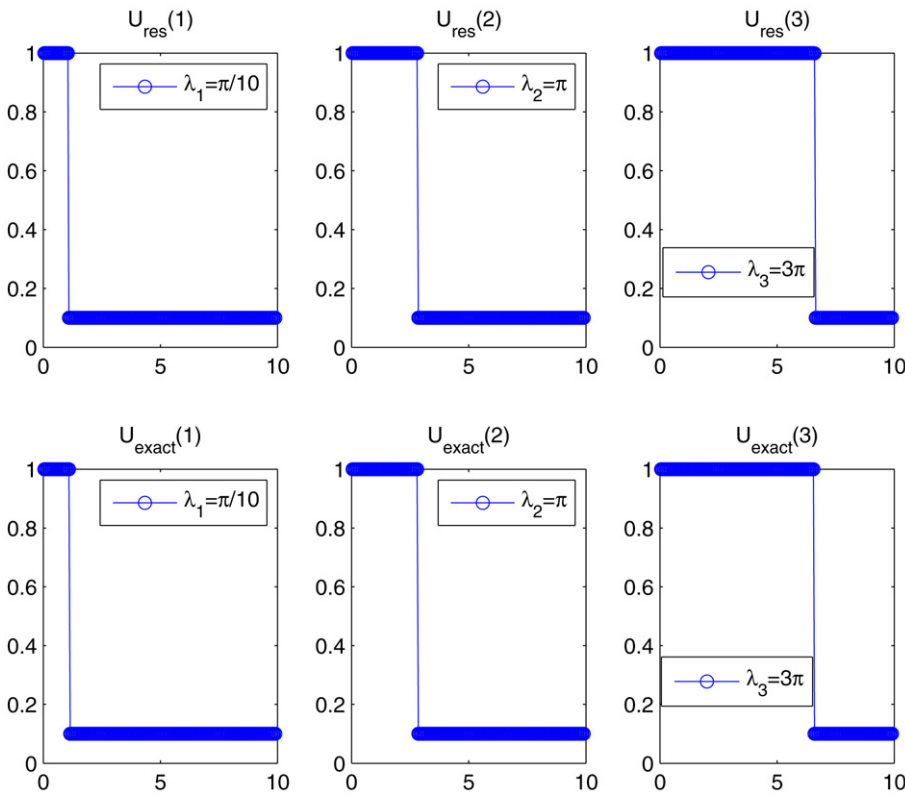


Fig. 9. Reservoir and exact solutions at time  $T = 0.6$  (irrational eigenvalues).

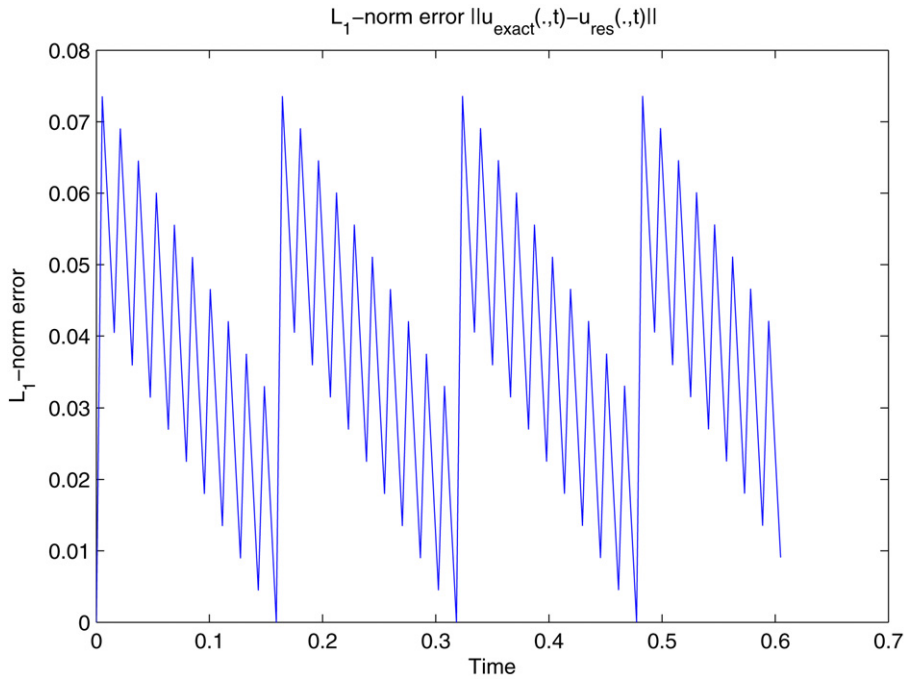


Fig. 10. Reservoir –  $L^1$ -norm error as a function of time (irrational eigenvalues).

systems, this method has indeed shown impressive results compared to classical high order methods such as ENO, WENO methods, etc., see [3]. The extension of Theorem 2.2 will be proven in the second part of this work.

**Theorem 3.1.** *The reservoir-Colella–Glaz scheme for hyperbolic systems of conservation laws is a convergent scheme and is exact at the discrete level for some general classes of initial data.*

Note that from Section 2, we can directly state that the above result is true for shock wave solutions (on each characteristic field) of nonlinear hyperbolic systems of conservation laws.

## References

- [1] F. Alouges, F. De Vuyst, G. Le Coq, E. Lorin, The reservoir scheme for systems of conservation laws, in: *Finite Volumes for Complex Applications, III*, Porquerolles, 2002, Lab. Anal. Topol. Probab. CNRS, Marseille, 2002, pp. 247–254 (electronic).
- [2] F. Alouges, F. De Vuyst, G. Le Coq, E. Lorin, Un procédé de réduction de la diffusion numérique des schémas à différence de flux d'ordre un pour les systèmes hyperboliques non linéaires, *C. R. Math. Acad. Sci. Paris* 335 (7) (2002) 627–632.
- [3] F. Alouges, F. De Vuyst, G. Le Coq, E. Lorin, The reservoir technique: A way to make Godunov-type schemes zero or very low diffusive. Application to Colella–Glaz, *Eur. J. Mech. B Fluids* 27 (6) (2008).
- [4] F. Alouges, G. Le Coq, E. Lorin, Two-dimensional extension of the reservoir technique for some linear advection equations, *J. Sci. Comput.* 31 (3) (2007).
- [5] S.J. Billett, E.F. Toro, On WAF-type schemes for multidimensional hyperbolic conservation laws, *J. Comput. Phys.* 130 (1) (1997) 1–24.
- [6] P. Colella, H.M. Glaz, Efficient solution algorithms for the Riemann problem for real gases, *J. Comput. Phys.* 59 (2) (1985) 264–289.
- [7] B. Després, F. Lagoutière, Contact discontinuity capturing schemes for linear advection and compressible gas dynamics, *J. Sci. Comput.* 16 (4) (2001) 479–524, (2002).
- [8] J.-M. Ghidaglia, Numerical computation of 3D two phase flow by finite volumes methods using flux schemes, in: *Finite Volumes for Complex Applications II*, Hermes Sci. Publ., Paris, 1999, pp. 69–94.
- [9] J.-M. Ghidaglia, A. Kumbaro, G. Le Coq, Une méthode “volumes finis” à flux caractéristiques pour la résolution numérique des systèmes hyperboliques de lois de conservation, *C. R. Acad. Sci. Paris Sér. I Math.* 322 (10) (1996) 981–988.
- [10] E. Godlewski, P.-A. Raviart, *Hyperbolic Systems of Conservation Laws, Mathématiques & Applications (Mathematics and Applications)*, vol. 3/4, Ellipses, Paris, 1991.
- [11] S.K. Godunov, A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics, *Mat. Sb. (N.S.)* 47 (89) (1959) 271–306.
- [12] R.J. LeVeque, H.C. Yee, A study of numerical methods for hyperbolic conservation laws with stiff source terms, *J. Comput. Phys.* 86 (1) (1990) 187–210.
- [13] P.L. Roe, Approximate Riemann solvers, parameter vectors, and difference schemes, *J. Comput. Phys.* 43 (2) (1981) 357–372.
- [14] C.-W. Shu, S. Osher, Efficient implementation of essentially nonoscillatory shock-capturing schemes, *J. Comput. Phys.* 77 (2) (1988) 439–471.